A Pilot Sampling Design for Estimating Outdoor Recreation Site Visits on the National Forests

S. J. Zarnoch, S.M. Kocis, H.K. Cordell, and D.B.K. English

The Authors

S.J. Zarnoch, Mathematical Statistician, USDA Forest Service, Southern Research Station, Asheville, NC; **S.M. Kocis,** Recreation Planner, USDA Forest Service, Washington Office (detached), Star Route Box 300, Bridgeville, CA; **H.K. Cordell,** Project Leader and **D.B.K. English,** Research Social Scientist, USDA Forest Service, Southern Research Station, Athens, GA, respectively.

A Pilot Sampling Design for Estimating Outdoor Recreation Site Visits on the National Forests

S.J. Zarnoch, S.M. Kocis, H.K. Cordell, and D.B.K. English

Abstract

A pilot sampling design is described for estimating site visits to National Forest System lands. The three-stage sampling design consisted of national forest ranger districts, site days within ranger districts, and last-exiting recreation visitors within site days. Stratification was used at both the primary and secondary stages. Ranger districts were stratified based on Bailey's ecoregions, while site days were stratified based on site type, season, and day type. Statistical methodology is presented to derive site-visit estimates at the site day, ranger district, and national levels. Results are presented to illustrate the magnitude of the site-visit estimates, their variability, and confidence intervals. With such information, an evaluation of the stratification variables is presented using the design effect and the relative hypothetical efficiency. Sample size analysis is performed to provide recommendations for future sample surveys to meet specified levels of precision.

Keywords: National forests, outdoor recreation, sampling, site visits.

Introduction

Outdoor recreation is becoming increasingly important on the national forests of the United States. In many areas, forest managers are giving higher priority to recreation resources than to timber harvests. Although the U.S. Department of Agriculture Forest Service (USDA Forest Service) has conducted periodic inventories of forest timber resources for over 70 years, the Agency neither has developed nor used a statistically based, nationally applied analog for estimating the volume of recreation use on national forests. In large measure, this is due to the complexity of sampling issues. However, dependable, accurate recreation-use estimates are important for national-, regional- and forest-level decisionmaking and planning. Specifically, they arc needed to determine benefits that accrue from recreational use, as well as the impacts of such use on other forest resources and on local economies. Moreover, use estimates are needed to identify trends in outdoor recreation and assess the effectiveness of Federal government programs.

Traditionally, national forest recreation-use estimates were developed from individual ranger district estimates, which then were aggregated upward to produce national estimates. In 1965, the Agency established a reporting system called Recreation Inventory Measurement. It required site-level reporting of recreation visits. A study conducted by Kocis (1986) showed that the ranger district estimates were derived by numerous and diverse methods, producing results that

ranged from absolute knowledge to little more than guesses. Other national recreation surveys have been conducted by telephone, but none has included use information specific to the national forests. Many may be of questionable statistical validity because they usually were taken at selected locations and targeted only one or, at best, a few types of recreation users. Additionally, there may be unresolved problems of nonrespondent bias and sample selection. Moreover, none has given much attention to the variability of such estimates, which is necessary to quantify precision and reliability. Gregoire and Buhyoff (1999) presented statistically based methodologies for estimating recreation use, and those methodologies should provide a good basis for future surveys.

In an attempt to obtain statistically valid estimates and develop a sampling methodology for monitoring levels of outdoor recreation, the National Recreation Use Pilot Study was formulated by the USDA Forest Service in 1996. Its major objectives were to develop and execute a national pilot sample survey designed to estimate the total annual number of recreation site visits to the national forests and to provide information for developing a more efficient sampling design. It included the evaluation of stratification variables, estimates of strata variances, and sample size recommendations to achieve required precision levels. In addition, information was collected on visitors (age, gender, and geographic origin) and characteristics of their visit (where and when).

Our objectives were to: (1) describe the stratified three-stage sampling design that was used; (2) present average daily site-visit and total annual site-visit estimates, sample variances, and confidence intervals at the ranger district and national level; (3) evaluate the stratification variables; and (4) determine appropriate sample size recommendations. We have focused on the statistical aspects of estimating site visits but have not included results or discussions about visit and visitor characteristics.

Methodology

The National Forest System

The National Forest System lands comprise approximately 192 million acres across the continental United States,

Alaska, and Puerto Rico. Administratively, the USDA Forest Service is composed of nine regional offices that constitute numerous national forests, each of which may have several ranger districts. Within any USDA Forest Service region, or even within a national forest, there may be a great variety of forest types and, therefore, any number of outdoor recreation opportunities. This could lead to large variability in recreation-use estimates within a region. Although regions and national forests may be potentially convenient administrative units, they may not form homogeneous groups of ranger districts necessary for effective stratification in a statistical estimation process. An alternative is to group ranger districts by the environmental characteristics described in Bailey's classification of ecoregions (Bailey 1995, Bailey and others 1994), which divides the continental United States into homogeneous regions based on forest type, geology, and weather patterns. To the extent that the amount and type of recreation is determined by these factors, ranger districts within each ecoregion may have similar recreation-use patterns and, consequently, less variation in recreation-use estimates. Therefore, we used the ecoregion as a stratification variable in our sample design.

The Site Visit

An essential element of any sampling design is precise definition of what is to be measured or estimated. We define a site visit as one individual traveling to a recreation site in a national forest for a variable length of stay for the purpose of recreating, and then departing. For example, an individual camping 1 week at a campground constitutes one site visit, while a father and son fishing one Saturday on a river constitutes two site visits. Nonetheless, although some national forest lands have discrete individual sites, others do not. Therefore, a clear definition of sites is required. Personnel from the ranger districts selected for our study were asked to categorize each site or area in their jurisdiction into one of five mutually exclusive site types. These were used as stratification variables in an attempt to reduce variation.

Day-Use Developed Sites (DUDS)-are intended for day use only and include boating areas, picnic sites, fish-viewing sites, fishing sites, information sites, interpretive sites, observation sites, playground-park sport sites, ski areas (alpine and Nordic), wildlife viewing areas, visitor centers, museums, swimming areas, and winter sport sites. Generally, DUDS provide visitor comfort, convenience, and educational opportunities, but they are available only on a day-use basis.

Overnight-Use Developed Sites (OUDS)—include campgrounds, cabins, hotels, lodges, resorts, horse camps, organization sites, and any other overnight facility on national forest lands, whether they are owned and/or managed by the USDA Forest Service or are a private concession.

Water-Based Area Sites (WBAS)—are used exclusively for water activities. Other uses normally associated with DUDS may occur occasionally, but are of only minor consequence. In addition, the comfort and convenience usually associated with DUDS are minimal or absent.

General Roaded Dispersed Sites (GRDS)—include forest areas not included in DUDS, OUDS, or WBAS that are accessible by roads.

Unroaded Dispersed Sites (URDS)—include forest areas not included in DUDS, OUDS, WBAS, or GRDS, but which are not accessible by road.

The GRDS and URDS are associated with hiking, hunting, and dispersed camping. They were created by mapping watersheds or land areas accessed by major roads or trails. Originally, we wanted to create GRDS and URDS of 2,000 to 5,000 acres; but, due to differences among ranger districts, size actually varied from 1,000 to 100,000 acres. Thus, each ranger district potentially had many GRDS and URDS, and each was treated as an individual site in the same manner as were DUDS, OUDS, and WBAS.

Site Days

Generally, the most basic component of a sample survey is the population of sampling units from which a sample is drawn according to the criteria of a particular sampling design. In order to estimate the number of annual site visits, let sampling variable y be the number of last-exiting recreationists on a given site day. Last-exiting recreationists are those who are leaving the site for the last time and will not return during the site visit. This is in contrast to recreationists who merely are leaving the site for a particular reason, e.g., shopping, sightseeing, but will return later during their site visit. For our purposes, a site day is defined as the 24-hour period in which a site is open for recreation. Thus if a site is open throughout the entire year, it represents 365 site days. If it is open only on weekends, it has 104 site days. The total collection of y units over all sites days in all ranger districts is the population of sampling units, i.e., the sum of the y's is the true total number of site visits. Hence, the sampling problem was to design an efficient method of sampling the population of site days over the entire National Forest System.

It should be noted that instead of y being the number of last-exiting recreationists, it just as easily could have been first-entering recreationists. If an estimate of site visits is all that is required of a survey, it makes no difference whether sampling is at the time of first entry or last exit. However, if any ancillary sampling variables associated with the visit are desired, as was true in our study, exit interviewing is preferred for several reasons. If interviewed at the start of the visit, individuals can only give predictions about what they expect from the visit. Exit interviews capture more precise information especially about such things as length of stay, facilities used, and recreation activities. Moreover, entry interviews may directly affect the recreation visit. Experience also has shown that many visitors are eager to begin their recreation visit and are less willing to be interviewed as they begin their visit. For these reasons we designed our survey to count and interview visitors on their last exit.

Sampling Design

We used a three-stage sampling design to estimate the total number of site visits. The population of primary sampling units comprised all 606 ranger districts in the National Forest System. The secondary sampling unit was the collection of all site days (the 24-hour period during which a recreation site was open for visitation). The tertiary sampling unit was the last-exiting recreation visitors sampled and interviewed using a short, 2-minute questionnaire.

The selection of sampling units for the three-stage sampling design could follow a variety of methods, e.g., random, stratified, and systematic. However, in most situations stratification is advantageous for numerous reasons (Cochran 1977, Kish 1965). Stratification can increase precision of estimates if it is possible to divide a heterogeneous population into strata that are internally more homogeneous. In addition, if the strata represent meaningful subdivisions of the population, estimates can be obtained for each stratum. Often it is administratively more convenient to use stratification because it will ensure the sample not only is spread over the whole population but also is divided into manageable subpopulations that may be sampled locally. Finally, since the most appropriate sampling methodology may differ across the population, stratification allows for different sampling designs among strata.

In an attempt to reduce variation, we stratified the primary sampling units (ranger districts) into 16 ecoregion strata based on Bailey's ecoregion classification (Bailey 1995, Bailey and others 1994). Each stratum is composed of entire ccoregion divisions or parts thereof, delineated to form a contiguous piece of land. This resulted in 16 ecoregion

strata formed in 10 of Bailey's 14 ecoregion divisions. Table 1 defines the strata by Bailey's ecoregion division and indicates the specific ecoregion provinces that they contain. We felt that this would result in more homogeneous strata than if arbitrary administrative units, such as national forests, were used. Due to limited budgets, we could sample only 32 ranger districts over all ecoregion strata, resulting in a 5.3-percent sample. We used proportional allocation to determine the number of sampled ranger districts per ecoregion stratum, stipulating that each had at least one ranger district. The original selection of ranger districts was random; however, some selected ranger districts could not participate. To ensure the appropriate sample size, we replaced them with other randomly selected ranger districts. Table 1 shows the number of ranger districts in each ecoregion stratum' and the number sampled based on approximate proportional allocation. During our study, the two ranger districts in ecoregion 7 dropped out, resulting in sampling of 30 ranger districts in 15 ecoregions.

The unique and diverse character of recreation sites, as well as the variety of and large expected differences in the volume of activities during different days and seasons, allowed us to stratify the secondary sampling unit (site day). Those strata were formed by the five site types, along with strata defined by two seasons and two day types. Season and day-type stratification variables were included to reduce variation caused by seasonal and daily fluctuations in site visits. We classified each site day within a given site type as belonging to either high or low seasonal use, not necessarily coinciding with the calendar year four seasons. The day-type strata classified site days as either weekdays or weekend/holidays. Thus, there were $5 \times 2 \times 2 = 20$ possible strata for the site days of each ranger district. We anticipated that these strata had the potential to classify site days into homogeneous groups that would reduce variance of the estimates. Available funding allowed us to sample approximately 70 site days per ranger district. The actual number of site days sampled for each ranger district, disregarding those that were missed for various reasons, e.g., dangerous weather conditions, personnel issues, is shown in table 2. Proportional allocation was used to distribute the sample of site days over the 20 potential strata within a ranger district, after allocating 2 site days per stratum.

The tertiary sampling unit was the recreation visitor, which was sampled by selecting vehicles exiting for the last time. We conducted a 2-minute interview while allowing other vehicles to exit and used a 24-hour vehicle counter to tally

---

[1] Ecoregion stratum will be referred to as simply ecoregion in the remainder of this manuscript.

Table 1-Ecoregion strata, Bailey's division and Bailey's province, total number of ranger districts, and number of ranger districts sampled

| Ecoregion stratum | Bailey's ecoregion | | Ranger districts | |
	Division	Province	Total	Sampled
			---- *Number* ----	
1	Marine (240)	M244, M245	14	1
2	Marine (240)	242, M242	57	3
3	Mediterranean (260)	26 1,262, M26 1, M262	82	4
4	Temperate desert (340)	341,342, M341	26	1
5	Tropical/subtropical steppe (3 10)	313		
	Tropical/subtropical desert (320)	321,322		
	Temperate steppe (330)	331	11	1
6	Temperate steppe (330)	331,332, M331, M333	107	4
7	Tropical/subtropical steppe (3 10)	M313		
	Temperate steppe (330)	M331, M334		
	Temperate desert (340)	341, M341, 342	28	2"
8	Tropical/subtropical steppe (3 10)	311,315		
	Temperate steppe (330)	331,332	60	3
9	Tropical/subtropical steppe (3 10)	M313	35	2
10	Prairie (250)	251		
	Temperate steppe (330)	331,332	19	1
11	Warm continental (2 IO)	212	37	2
12	Hot continental (220)	222	12	1
13	Hot continental (220)	M222		
	Subtropical (230)	231, M231	29	2
14	Hot continental (220)	221		
	Subtropical (230)	23 1, 232, 234	35	2
15	Warm continental (2 10)	212, M212		
	Hot continental (220)	222	13	1
16	Hot continental (220)	M221	4 1	2
Total			606	32

[a] These two ranger districts dropped out of the survey

Table 2—Individual ranger district average daily site visit and total annual site-visit estimates, standard errors, and coefficients of variation

Ecoregion stratum	Site days sampled'	Vehicles		Total site days'	Average daily site visits	SE	Total annual site visits	SE	CV
		Stopped	Interviewed [b]						
				- - - - - - - - - - - - - *Number* - - - - - - - - - - - - -					
1	80	471	177	4,878	18.6	5.36	90,792	26,151	28.8
2	74	531	420	8,239	144.4	24.49	1,194,108	201,736	17.0
	50	360	234	7,070	81.7	35.55	577,899	251,370	43.5
	55	219	107	8,073	16.5	4.10	133,265	33,120	24.9
3	6.5	526	394	13,331	123.0	47.41	1,639,331	631,986	38.6
	65	261	135	13,118	46.8	12.45	614,503	163,257	26.6
	74	313	129	6,606	31.1	7.82	205,507	51,683	25.1
	13	43	16	3,558	6.1	2.85	21,527	10,158	47.2
4	98	566	392	9,573	70.6	43.18	675,629	413,376	61.2
5	77	797	647	14,947	101.1	21.09	1,510,905	315,222	20.9
6	60	130	71	5,450	12.6	4.28	68,794	23,311	33.9
	72	111	69	4,785	45.6	30.50	218,344	145,964	66.9
	60	109	86	5,641	10.5	2.23	59,321	12,597	21.2
	48	180	94	2,912	37.2	11.50	108,244	33,482	30.9
8	47	600	507	21,784	22.5	3.40	490,882	74,129	15.1
	74	616	437	6,590	46.3	9.80	305,283	64,573	21.2
	69	652	509	12,182	88.9	20.22	1,082,687	246,290	22.7
9	81	1,628	1,444	10,487	208.5	43.48	2,186,175	455,974	20.9
	85	508	334	8,164	118.8	24.40	969,482	199,236	20.6
10	70	157	97	6,560	14.1	2.73	92,719	17,896	19.3
11	85	229	126	8,396	16.9	4.28	142,112	35,916	25.3
	44	342	161	3,831	92.9	23.64	355,738	90,575	25.5
12	7.5	483	268	7,726	67.1	6.95	5 18,283	53,664	10.4
13	79	249	55	11,184	11.6	2.59	130,284	29,010	22.3
	66	387	242	6,593	27.0	8.92	178,260	58,832	33.0
14	82	473	302	7,923	29.4	7.12	262,986	56,419	24.2
	86	1,079	447	9,189	63.1	14.32	579,565	131,583	22.7
15	80	309	219	7,921	13.6	3.64	107,653	28,824	26.8
16	58	451	334	11,682	125.3	37.20	1,463,958	434,617	29.7
	59	86	32	7,782	13.8	7.41	107,090	57,649	53.8

SE = standard errors; CV = coefficients of variation.

[a] Site days sampled per ranger district.

[b] Represents vehicles the occupants of which had been recreating, were willing to be interviewed, and were exiting the site for the last time.

[c] Total site day is the total annual number of site days on the ranger district.

all exiting vehicles. The interviews served to calibrate the 24-hour vehicle count to produce an estimate of the sampling variable y as described in the next section. The total number of vehicles stopped for a potential complete interview and the total number of last-exiting recreating vehicles whose occupants were willing to complete the interview are shown in table 2.

There were several advantages to using the three-stage sampling approach over simple random sampling. Transportation costs were reduced because sampling was clustered within several primary sampling units and not spread out over the entire population of ranger districts. Logistics and administrative details were simplified because sampling was made on far fewer ranger districts. By narrowing our sample of ranger districts, we were able to interact with and train personnel conducting the field sampling.

Site-Visit Estimates for a Site Day

We used a double sampling technique (James 1967, James and Henley 1968, James and Ripley 1963) to estimate the number of recreation site visits per site day. Using a one-way exit-vehicle counter, we recorded the number of vehicles exiting a recreation site during a 24-hour period. To calibrate the 24-hour vehicle count, we interviewed a random sample of last-exiting visitors on the same day. An estimate of site visits per day was defined as:

$$y^{(1)} = LEV\ \overline{X} , \qquad (1)$$

where

$y^{(1)} =$ an estimate of the number of site visits on a given site day,
$LEV =$ the number of recreating vehicles exiting the site for the last time during the 24-hour period, and
$\overline{X} =$ the average number of people in a recreating vehicle exiting the site for the last time.

We administered a questionnaire to a random sample of all vehicles exiting the site during the interviewing period, which usually lasted about 6 hours. The initial questions screened out all but the last-exiting recreating vehicles, to which a more intensive set of questions was given. This produced information about duration of site visit, recreation activities, citizenship, gender and age class, and mode of transportation. However, our study focused on site visits, not the characteristics of visits or visitors. From the interview data, we calculated the average number of people in a vehicle, $\overline{X}$, based only on those vehicles that were recreating and exiting the site for the last time.

To obtain an estimate of *LEV,* we used both sampling devises. Because *LEV* was based on last-exiting recreating vehicles, that number could not be determined by a vehicle counter that could not distinguish nonrecreating vehicles or recreating vehicles that were not exiting for the last time. Therefore, we used information from the questionnaire to estimate the proportion, *P,* of all exiting vehicles whose occupants had been recreating and were exiting for the last time. We assumed that *P* estimated from interviews would provide a reasonable estimate of the proportion of last-exiting recreating vehicles in a 24-hour period.

One-way counters are preferred over two-way counters. If a two-way counter is used, as it was at some of the ranger districts, a one-way exit vehicle count can be approximated by dividing the count by two. Such an adjustment assumes that vehicles enter and exit at the same rate. If this assumption is faulty, a biased estimate of exiting traffic results.

An estimate of *LEV* was then defined as:

$$LEV = C * P , \qquad (2)$$

where

$LEV =$ the number of last-exiting recreating vehicles,
$C =$ the number of exits recorded by the vehicle counter during the 24-hour period, and
$P =$ the proportion of exiting vehicles that were last-exiting recreating vehicles.
To accommodate for variations in vehicle counters among the ranger districts, C was calculated as:

$$C = \frac{END - BEGIN}{WAY * INTERVAL * AXLES} , \qquad (3)$$

where

$BEGIN =$ the beginning vehicle count,
$END =$ the ending vehicle count,
$WAY =$ the number of traffic directions (one way or two way) in which the vehicle counter operated,
$INTERVAL =$ the interval length (proportion of a 24-hour day) in which the vehicle counter was operable, and
$AXLES =$ the average number of axles per vehicle estimated from interviews (this is set to 1 when vehicle counters do not count axles but count vehicles instead).

Although the daily site-visit estimator $y^{(1)}$ based on interviews and vehicle counters was preferred, difficulties prohibiting its use sometimes occurred. For example, it could

not be used when no interviews had been conducted, or when the vehicle counter did not operate. Alternative site-visit estimators were defined as:

$y^{(2)}$ = the total number of people counted by the interviewer (this should be biased low),
$y^{(3)}$ = the total number of people counted by the interviewer expanded to a 16-hour day, which represents the typical time period when recreationists might be exiting the site, and
$y^{(4)}$ = an alternative, reliable source such as ticket receipt counts (a nonestimated, observable count).

If, on occasion, information based on an alternative, reliable source was available, such as ski rental ticket receipts or park entrance fee receipts, then $y^{(4)}$ was used instead of $y^{(1)}$. If no vehicle count was available, then $y^{(3)}$ was used. If both interviews and vehicle count were available, then $y^{(1)}$ was used, unless $y^{(1)}$ was less than $y^{(2)}$, in which case $y^{(2)}$ was the site-visit estimate. In situations where no site-visit estimator was available, e.g., no one was interviewed, no vehicles were counted, and no alternative, reliable sources were available, then it was simply assigned zero.

Site-Visit Estimates for all Site Days on a Ranger District

Site-visit estimates for individual ranger districts were obtained using a stratified random sampling design where 20 strata consisted of the 5 site types, 2 seasons, and 2-day types (see Sampling Design). Let:

N = the total annual number of site days on the ranger district,
L = the total number of strata on the ranger district,
N_h = the total annual number of site days on the ranger district for stratum h (h = 1, 2, 3, . . . , 20),
$W_h = N_h /N$ = stratum h weight,
n = the total number of site days sampled in all strata,
n_h = the number of site days sampled in stratum h, and
y_{hi} = the site-visit estimate in stratum h for day i.

Note that y_{hi} may be based on the $y^{(1)}$, $y^{(2)}$, $y^{(3)}$, or $y^{(4)}$ site-visit estimators, depending on the data available for the specific site day as discussed in the previous subsection. Then the average daily site-visit estimate in stratum h is defined as:

$$\bar{y}_h = \sum_{i=1}^{n_h} \frac{y_{hi}}{n_h}, \qquad (4)$$

and the estimated variance in stratum h is:

$$s_h^2 = \sum_{i=1}^{n_h} \frac{\left(y_{hi} - \bar{y}_h\right)^2}{\left(n_h - 1\right)}. \qquad (5)$$

The overall average daily site-visit estimate for the entire ranger district is:

$$\bar{y} = \sum_{h=1}^{L} W_h \, \bar{y}_h, \qquad (6)$$

with estimated variance:

$$s^2\left(\bar{y}\right) = \sum_{h=1}^{L} \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^{L} \frac{W_h s_h^2}{N}. \qquad (7)$$

Approximate confidence intervals may be calculated in the typical manner as:

$$\bar{y} \pm z \, s\left(\bar{y}\right), \qquad (8)$$

where

z = the z-value from the standard normal distribution at the appropriate percentage point (Montgomery 1976), for example, z = 1.96 at the 95-percent confidence level.

An estimate of the total annual number of site visits for a ranger district is easily obtained by expanding the average daily site-visit estimate [equation (6)] by the total annual number of site days on the ranger district. Mathematically, this is defined as:

$$Y = N \bar{y} = \sum_{h=1}^{L} N_h \, \bar{y}_h, \qquad (9)$$

with estimated variance:

$$s^2\left(Y\right) = N^2 s^2\left(\bar{y}\right). \qquad (10)$$

An approximate confidence interval would be:

$$Y \pm z \, s\left(Y\right) \qquad (11)$$

Two types of sampling allocation are commonly used with stratified random sampling-proportional or optimum. Our study used proportional allocation, which is defined as:

$$n_h = n \, W_h. \qquad (12)$$

It assigns the number of sampling units to a stratum in proportion to its size. Optimum allocation, used later in the evaluation of the stratification variables, is defined as:

$$n_h = n \frac{W_h s_h}{\sum_{i=1}^{L} W_h s_h}. \qquad (13)$$

This allocation method assigns a larger sample to a stratum if the stratum is larger or is more variable internally.

Site-Visit Estimates for the Entire National Forest System

The site-visit estimate for the entire National Forest System was obtained by estimating site visits at the ecoregion level and expanding them to the national level. Let

N_i = the number of ranger districts in ecoregion i,

then the average total annual site-visit estimate for a ranger district in ecoregion i is:

$$\bar{Y}_i = \frac{\sum_{j=1}^{n_i} Y_{ij}}{n_i}, \qquad (14)$$

where

I', = the total annual site-visit estimate for ranger district j in ecoregion i, and
n_i = the number of ranger districts sampled in ecoregion i.

Theoretically, the sample variance of $\bar{Y}_i$ consists of two components-the variances between and within ranger districts. However, because the number of ranger districts within each ecoregion was low, i.e., n, IN, is negligible, the variance within ranger districts could be ignored as explained by Cochran (1977, p. 279), yielding the sample variance as:

$$s^2\left(\bar{Y}_i\right) = \frac{\sum_{i=1}^{n_i}\left(Y_{ij} - \bar{Y}_i\right)}{n_i\left(n_i - 1\right)}\left(1 - \frac{n_i}{N_i}\right). \qquad (15)$$

The grand total annual site-visit estimate over all ecoregions in the National Forest System is obtained by expansion as:

$$SV = \sum_{i=1}^{16} N_i \bar{Y}_i, \qquad (16)$$

with estimated variance:

$$s^2\left(SV\right) = \sum_{i=1}^{16} N_i^2 \, s^2\left(\bar{Y}_i\right). \qquad (17)$$

Evaluation of Stratification

One way to evaluate the effect of stratification is to compare the variance of a simple random sample, V_{ran}, and the variance of a stratified random sample under optimum allocation, V_{opt} Cochran (1977) shows that:

$$V_{opt} = V_{ran} - \frac{1}{n}\sum_{h=1}^{L} W_h\left(S_h - \bar{S}\right)^2 - \frac{(1-f)}{n}\sum_{h=1}^{L} W_h\left(\bar{Y}_h - \bar{Y}\right)^2, \qquad (18)$$

where

$\bar{Y}_h$ = the true stratum h mean,
f = the sampling fraction, n/N,
$\bar{Y}$ = the true population mean,

S_h = the true stratum h standard deviation,

$\bar{S} = \sum_{h=1}^{L} W_h S_h$ = the true population standard deviation, and the other terms are as previously defined.

The term on the far right reflects the gain due to strata that have different means. The middle term to the right of the equal sign is the difference in variance between optimum and proportional allocation and is associated with differences in the stratum standard deviations. Therefore, the criteria used for evaluating the effectiveness of stratified random sampling under optimum allocation are the stratum means and standard deviations. Stratum differences in either of these indicate that stratification is effective. The former could be tested with a fixed-effects analysis of variance on the stratum means (Montgomery 1976, SAS Institute Inc. 1989). In simple situations, such as where only one stratification variable is being studied, the standard deviations could be tested with Bartlett's test for homogeneity of variances (Milliken and Johnson 1984, Montgomery 1976). However, we used the fixed-effects analysis of variance to test for variance differences because multiple stratification variables were being considered in a factorial fashion. Because variances usually exhibit a positively skewed distribution, they were converted to rank statistics to better satisfy the normality assumption required for the analysis of variance, resulting in a nonparametric statistical test (Conover 1980, SAS Institute Inc. 1990).

The design effect (DEFF) is another method used to compare the efficiency of a stratified random sampling design to a simple random sample (Cochran 1977). The DEFF is the ratio of the stratified random sampling variance of the estimate to the variance of the estimate from a simple random sample of the same sample size and, thus, quantifies the gain due to stratification. When DEFF = 1, the sampling methods are, equivalent. However, if DEFF is < 1, the stratified random sampling design has reduced the estimate's variance. Stratified random sampling is somewhat more complex and costly than simple random sampling; thus, DEFF must be substantially < 1 if stratification is to be of practical value. In the unlikely event it is > 1, simple random sampling is more precise. Rao (1962) defines an unbiased estimator of the variance of the mean of a simple random sample in terms of a stratified random sample as:

$$v'\left(\bar{y}\right) = \frac{(N-n)}{n(N-1)}\left(\frac{1}{N}\sum_{h=1}^{L}\frac{N_h}{n_h}\sum_{i=1}^{n_h} y_{hi}^2 - \bar{y}_{st}^2 + s^2\left(\bar{y}_{st}\right)\right), \qquad (19)$$

where

$\bar{y}_{st}$ = the mean estimate based on the stratified random sample,

$s^2\left(\overline{y}_{st}\right)=$ the variance based on the stratified random sample,

and the other terms are as previously defined.

The DEFF is then:

$$DEFF = \frac{s^2\left(\overline{y}_{st}\right)}{v'\left(\overline{y}\right)}. \qquad (20)$$

If the stratified random sampling design was performed under suboptimum allocation, where the n_h 's differed substantially from what they would be under optimal allocation, the DEFF may be above 1 even if stratification has potential. Thus, the evaluation of the stratification variables can be hypothetically evaluated under optimal allocation even though our study deviated from optimal allocation. The criterion is the relative hypothetical efficiency (RHE) (Ruark and Zamoch 1992, Zamoch and others 1993) defined as:

$$RHE = \frac{v'\left(\overline{y}_{st}\right)}{v'\left(\overline{y}\right)}, \qquad (21)$$

where $v'\left(\overline{y}_{st}\right)$ is the estimated hypothetical variance based on the variance estimates s_h^2 when the stratum sizes, n_h, are determined by optimum allocation [equation (13)]. The DEFF will decrease to < 1 as the strata means become increasingly different. However, this is not necessarily true for differences in the stratum variances, because the DEFF will be approximately 1 under proportional allocation and will only diminish as optimal allocation is approached. Thus, RHE is useful in detecting effective stratification due to differences in variances when optimum allocation is used.

Results and Discussion

Site-Visit Estimates for Individual Ranger Districts

Total number of site days per year per ranger district varied from 2,912 to 21,784 and averaged 8,539. In general, this should be highly correlated with not only the number of recreational sites and number of days the recreational sites are open but also with how the sites are defined in the field. This is particularly true of GRDS and URDS sites that were quite arbitrarily defined. The number of site days selected for sampling per ranger district varied from 62 to 98; however, due to various factors such as bad weather, personnel issues, and emergencies, the number actually sampled ranged from 13 to 98 (average 68).

Estimates for average daily site visits [equation (6)] and total annual site visits [equation (9)] for each ranger district are

shown in table 2. The average daily site-visit estimate was 56.9 people and varied across ranger districts from 6.1 to 208.5. Total annual site visits per ranger district averaged 536,378, ranging from 21,527 to 2,186,175.

The precision of these estimates varied considerably and is reflected in the standard errors. However, it is often more informative to evaluate precision with the coefficient of variation, which puts the standard error on a relative basis to the parameter that is being estimated. The coefficient of variation ranged from 10.4, which is considered quite good in this type of survey, to a high of 66.9 (average 29.3). It is also useful to construct confidence intervals [equations (8) and (11)] from which the precision of the estimate can be appreciated even more easily. For example, a 90-percent confidence interval on the total annual site-visit estimate for the most precise ranger district (ecoregion 12 in table 2) is 430,006 to 606,560, which implies a considerably tight bound on the estimate. However, when a 90-percent confidence interval is constructed for the most variable ranger district (second ranger district in ecoregion 6 in table 2), the bounds are 0 to 458,454, indicating that this estimate is practically useless.

Site-Visit Estimate for Entire National Forest System

The total annual site-visit estimate for National Forest System lands was 330,772,894 [equation (16)] with a standard error of 66,280,755 [equation (17)], yielding a coefficient of variation of 20 percent. The 80-percent confidence interval was calculated to be 245,800,966 to 415,744,822, which indicates that the estimate was within 25.7 percent of the true number of site visits with a probability of 0.80. This is a reasonable level of precision for a study of this type. The goal for future recreation surveys is to be within 15 percent with a probability of 0.80.

The true cost of sampling is hard to determine because we used a variety of full-time paid employees and volunteers for sampling. We asked ranger district officials to estimate the survey's true cost. Estimated startup costs ranged from $300 to $2,387 and a sampled site-day cost from $58 to $240. Therefore, ranger district funding was about $200 per sampled site day, including $150 to set up a vehicle counter and conduct on-site interviews, and $50 for retrieving the counter the next day. Overhead and vehicle expenses were included in these costs but equipment purchases were not.

Calculating the national site-visit estimate included a slight modification, because both ranger districts in ecoregion 7 dropped out of the survey. Basically, a preliminary national

site-visit estimate and variance were calculated from 15 ecoregions comprising 578 ranger districts. The final national site-visit estimate was obtained by multiplying the preliminary estimate by the expansion factor 6061578 to reflect the unsampled ecoregion 7, which had 28 ranger districts. This assumes that the recreation-use level per ranger district in ecoregion 7 was not different from the overall national average. The variance was also expanded by multiplying by the square of the expansion factor.

Evaluation of Stratification

Overview-Our study used stratification to reduce variance of the estimates. Stratification was used at both stages of the sampling design: the primary units (ranger districts) were stratified based on 16 ecoregions, while the secondary units (site days) were stratified based on 5 site types, 2 seasons, and 2 day types. There were very few sampled ranger districts in each ecoregion (table 1). Because sample variances for each stratum require at least two observations, almost half of all ecoregions had no sample variance estimate, while others were based on an extremely low sample size. Only poor estimates of the stratum sample variances were obtainable; thus, no evaluation of the ecoregion strata was attempted.

Stratification at the secondary stage (site days) produced a better estimate of the sample variance in each stratum. However, even here small sample sizes hindered the estimates. Evaluation of stratification at this stage focused on determining the usefulness of the site-type, season, and day-type strata for reducing variation in site-visit estimates. The analysis of variance was used to test for differences between stratum means and variances for the site-day strata on an ecoregion basis. In addition, the DEFF and RHE criteria were also computed and compared in a similar manner.

Analysis of stratum means-Stratification effectiveness was assessed based on the relationship between simple random sampling and stratified random sampling under optimum allocation [equation (1 S)]. Differences between stratum means were tested using analysis of variance procedures for a randomized block design. The ranger districts were considered blocks, while the three stratification variables were site type, season, and day type. We conducted the analysis by ecoregion, and then pooled all ecoregions.

The analysis of variance initially was performed on the full model by including all two- and three-way interactions of the stratification variables. On an ecoregion basis, very few interactions were significant. Site type*season was only significant for ecoregions 2, 3, and 13; and site type*day

type and site type*season*day type were significant only for ecoregion 13. When pooled over ecoregions, site type*season was significant (p = 0.0223); however, site type*day type (p = 0.3754), season*day type (p = 0.5840), and site type*season*day type (p = 0.419 1) were not significant. Only main effects were considered further because interactions were mostly nonsignificant and those that were significant would be difficult to interpret.

Using a main effects model on an ecoregion basis, analysis of variance still showed very few statistical significances at the 0.05 level, even though the strata means appeared quite different (table 3). However, when all ranger districts were pooled across the ecoregions, all stratification variables showed statistical significance. Overall, trends within each of the stratification variables seemed reasonable, but they were not consistent among the ecoregions. For instance, within the site-type strata, DUDS had the highest overall mean site-visit estimate (106.8) while URDS had the lowest (37.2). However, this relationship only held up on a few of the ecoregions. In particular, in ecoregions 3, 10, 12, and 14, DUDS had the lowest mean site-visit estimate. Although this may be due to high variability, differences may also have been the result of different recreational patterns among the ecoregions. The seasonal strata means were different in a consistent manner across all ecoregions. High season was greater than low season in all cases except ecoregion 15. The day-type strata were not quite as consistent but generally exhibited the expected pattern, i.e., weekend/holiday days were greater than weekdays, although the weekend/holiday mean occasionally was less than the weekday mean. Thus, although statistical significances were for the most part lacking, site type and season tended to exhibit different strata means and, hence, appear as potentially feasible stratification candidates.

Analysis of stratum variances-Stratification effectiveness was further tested by using analysis of variance procedures similar to those used for the stratum means, except that the variable tested was the rank statistics of the stratum variances. Results from the full model, including interactions on an ecoregion basis, revealed that season*day type was significant only in ecoregions 2, 3, and I 1; and that site type*season and site type*season*day type were significant only in ecoregions 11 and 13. However, when pooled over ecoregions, site type*season (p = 0.4995) site type*day type (p = 0.3706) season*day type (p = 0. 1000), and site type*season*day type (p = 0.3968) were all not significant. Again, only main effects were considered further.

For the main effects model on an ecoregion basis, the analysis of variance showed five significances for site type, five for season, and only one for day type (table 4). When

Table 3--Analysis of variance on the stratum mean site-visit estimates for three stratification variables and their associated least squares means

	Analysis of variance			Site type					Season [a]		Day type [b]	
Ecoregion	Source	df	P-value	DUDS	GRDS	OUDS	URDS	WBAS	H	L	H	L
				• • • • • • • • • • • • • Least squares means **of** site-visit estimates • • • • • • • • • • • • •								
1	Site type	4	0.25	25.9	12.3	11.0	9.6	3.4				
	Season	1	.07						19.1	5.8		
	Day type	1	.53								14.5	10.5
2	Site type	4	.08	216.6	97.4	84.3	17.3	66.7				
	Season	1	.06						140.3	52.6		
	Day type	1	.76								103.4	89.5
3	Site type	4	.12	34.9	63.4	43.8	35.9	139.5				
	Season	1	.21						78.I	48.9		
	Day type	1	.84								65.8	61.2
4	Site type	4	.36	72.5	33.9	148.5	15.4	37.4				
	Season	1	.06						106.7	16.4		
	Day type	1	.88								58.1	64.9
5	Site type	4	.35	186.3	55.4	73.8	152.7	236.8				
	Season	1	.42						165.9	116.1		
	Day type	1	.12								188.1	93.9
6	Site type	4	.46	18.7	29.1	14.0	32.8	48.1				
	Season	1	.29						34.3	22.8		
	Day type	1	.18								35.8	21.3
8	Site type	4	.05	168.7	29.9	19.7	39.8	70.9				
	Season	1	.04						103.8	27.8		
	Day type	1	.84								62.1	69.5
9	Site type	4	.01	524.6	124.6	29.6	80.1	242.7				
	Season	1	.14						274.0	126.7		
	Day type	1	.75								215.5	185.1
10	Site type	3	.15	6.6	19.1	43.2	⸺	8.6				
	Season	1	.20						27.1	11.6		
	Day type	1	.27								26.0	12.8
11	Site type	4	.12	53.5	94.2	44.2	39.4	19.0				
	Season	1	.48						56.9	43.2		
	Day type	1	.21								62.1	38.0
12	Site type	4	.22	10.5	53.2	59.0	154.1	51.8				
	Season	1	.26						87.6	43.9		
	Day type	1	.13								94.3	37.2
13	Site type	4	.15	12.2	13.0	34.8	9.6	19.6				
	Season	1	.08						24.1	11.7		
	Day type	1	.03								25.6	10.2
14	Site type	4	.07	-3.6'	16.0	57.0	29.1	205.8				
	Season	1	.13						98.8	22.9		
	Day type	1	.17								92.9	28.8
15	Site type	4	.60	15.2	5.4	12.0	5.6	33.1				
	Season	1	.72						12.0	16.5		
	Day type	1	.22								6.4	22.1
16	Site type	4	.26	5.1	169.2	196.4	28.1	-43.9"				
	Season	1	.33						109.0	33.0		
	Day type	1	.23								117.4	24.6
All	Site type	4	.01	106.8	59.8	56.9	31.2	79.8				
	Season	1	0						92.3	43.9		
	Day type	1	.05								80.0	56.2

df = degrees of freedom; DUDS = day-use developed sites stratum; GRDS = general roaded dispersed sites stratum; OUDS = overnight-use developed sites stratum; URDS = unroaded dispersed sites stratum; WBAS = water-based area sites stratum.

[a] H = high season stratum; L = low season stratum.

[b] H = high (weekends/holidays) day-type stratum; L = low (weekdays) day-type stratum.

[c] Since least squares means are computed from a general linear model, negative values are possible.

Table 4-Analysis of variance on the stratum rank variances of the site-visit estimates for three stratification variables along with the stratum average variances (1,000s)

Ecoregion	Source	df	P-value	DUDS	GRDS	OUDS	URDS	WBAS	Season[a] H	Season[a] L	Day type[b] H	Day type[b] L
				▪ ▪ ▪ - - - - - - - - - - - *Estimated average of site-visit days* - - - - - - - - - - - - - ▪								
1	Site type	4	0.33	0.6	1.4	0.5	0.4	0				
	Season	1	.09						0.7	0.4		
	Day type	1	.48								0.5	0.6
2	Site type	4	.44	262.3	34.6	13.9	7.9	18.1				
	Season	1	.01						119.2	11.3		
	Day type	1	.50								19.1	126.1
3	Site type	4	.39	18.9	4.1	9.1	1.5	98.9				
	Season	1	.12						31.4	2.6		
	Day type	1	.66								4.3	40.0
4	Site type	4	.56	2.5	1.2	444.5	.6	2.8				
	Season	1	0						180.2	.3		
	Day type	1	.30								2.7	197.4
5	Site type	4	.04	68.6	.8	14.0	53.4	4.3				
	Season	1	.21						42.0	17.0		
	Day type	1	.05								36.6	25.2
6	Site type	4	.46	.7	3.6	.7	25.0	2.8				
	Season	1	.01						4.1	12.5		
	Day type	1	.94								4.3	11.1
8	Site type	4	0	70.8	2.1	.3	10.4	2.5				
	Season	1	.04						39.1	1.7		
	Day type	1	.63								21.3	19.8
9	Site type	4	.02	311.5	72.2	.4	17.5	143.1				
	Season	1	.62						164.4	63.3		
	Day type	1	.99								112.1	115.3
10	Site type	3	.13	.1	.9	6.3		.2				
	Season	1	.13						2.7	.2		
	Day type	1	.71								2.9	.4
11	Site type	4	.01	6.5	24.4	.3	.2	.7				
	Season	1	0						7.8	7.4		
	Day type	1	.25								9.0	6.3
12	Site type	4	.28	.7	3.2	2.6	6.3	1.2				
	Season	1	.98						3.1	3.3		
	Day type	1	.66								3.6	2.9
13	Site type	4	.05	.2	.8	4.6	.2	1.4				
	Season	1	.13						2.9	.3		
	Day type	1	.11								2.8	3
14	Site type	4	.09	1.1	4.2	12.1	2.8	93.1				
	Season	1	.40						30.2	7.5		
	Day type	1	.08								23.3	18.4
15	Site type	4	.67	.7	.2	.2	.2	6.3				
	Season	1	.81						.3	3.2		
	Day type	1	.22								.1	3.0
16	Site type	4	.06	1.1	68.8	596.6	3.7	2.1				
	Season	1	.26						223.8	15.2		
	Day type	1	.76								268.0	5.4
All	Site type	4	.01	75.7	17.2	54.7	10.2	26.3				
	Season	1	0						56.3	10.9		
	Day type	1	.09								32.7	39.1

df = degrees of freedom; DUDS = day-use developed sites stratum; GRDS = general roaded dispersed sites stratum; OUDS = overnight-use developed sites stratum; URDS = unroaded dispersed sites stratum; WBAS = water-based area sites stratum.
[a] H = high season stratum; L = low season stratum.
[b] H = high (weekends/holidays) day-type stratum; L = low (weekdays) day-type stratum.

pooled over all ecoregions, site type and season were significant; day type was not. There were no consistent trends for site type and day type among the ecoregions; however, season usually exhibited greater variability in the high stratum. Generally, similar to the stratum means analysis, these results indicate that site type and season are potentially feasible stratification candidates.

The DEFF and RHE analysis-An analysis of the DEFF [equation (20)] also was performed to assess strata usefulness (table 5). When all 20 strata were considered jointly, the mean DEFF was 1.01 On some ranger districts, **DEFF** values were close to 0.50, indicating that the strata were beneficial. However, caution must be used in interpreting DEFFs, because the stratified random sampling was performed under suboptimal allocation, i.e., not according to equation (13). In this situation the DEFF may be above 1 even if stratification has potential. To analyze the effect of stratification under optimum allocation, we used the **RHE** [equation (21)]. Our analysis showed that the average RHE was 0.40 and ranged from 0.12 to 0.75. Therefore, for all 30 ranger districts, the strata used have shown a significant reduction in variance if optimal allocation is used to assign sample sizes to the strata.

The DEFF and RHE were also analyzed on each of the three individual stratification variables separately. Each analysis on a stratification variable was performed by simply disregarding the other two strata variables (table 5). Site type clearly showed the most potential use for stratification because numerous DEFFs were between 0.50 and 0.75, while season and day type had none in that range. Also, the mean RHE for site type (0.64) was substantially lower than that for either season (0.84) or day type (0.89).

Sample size-ranger district level-Determining a recommended sample size for future surveys at the ranger district level requires obtaining good stratum variance estimates [equation (5)] from the pilot study, computing the appropriate variance formula [equations (7) and (IO)], and determining the appropriate sample size to meet a desired level of precision. However, because most strata had very few sample days, the reliability of stratum variance estimates for most individual ranger districts was poor. To increase the sample size upon which the stratum variances were based, we pooled ranger districts by ecoregions. Only ecoregions 2, 6, 13, and 14 had a complete set of stratum variances after pooling, so only these were used for the sample size analysis. In order to pool variances from different ranger districts within an ecoregion, we felt it was inappropriate simply to use an average stratum variance estimate [equation (5)] weighted by sample size for any given stratum because the ranger districts had greatly different levels of site

visitation. Consequently, stratum variances for each ranger district were converted into relative coefficients of variation (RCV) defined as:

$$RCV = 100 \frac{s_h}{\overline{Y}}, \tag{22}$$

where $\overline{y}$ and s_h are on a ranger district basis.

Computing the mean RCV for stratum h over all ranger districts in an ecoregion provides a relative measure of variation conditioned on the ranger district site-visitation level. The mean RCV for each stratum was squared and converted back to yield the s^2_h's for a given ranger district by multiplying by the ranger district's specific $(\overline{y}/100)$ factor. Thus, a set of stratum variances was obtained for a ranger district based on data pooled from all the ranger districts in one ecoregion. The mean RCV for the 20 strata for each of the 4 ecoregions reveal no obvious pattern within or between ecoregions (table 6). This is probably due to large sampling variability and relatively small sample sizes, even though ranger districts were pooled.

Sample size analysis was based on computing the variance of the average daily site-visit estimate for each of the ranger districts under a stratified random sampling design using optimum allocation. An estimate of the sample variance under optimum allocation is derived by substituting the n_h formula for optimum allocation [equation (13)] into the general sample variance formula [equation (7)], yielding:

$$s^2\left(\overline{y}_{opt}\right) = \frac{\left(\sum\limits_{h=1}^{L} W_h s_h\right)^2}{n} - \frac{\sum\limits_{h=1}^{L} W_h s_h^2}{N}, \tag{23}$$

where the W_h's and N are specific to the ranger district, and the s_h^2's are computed from the mean RCVs for the ecoregion (table 6).

To facilitate comparisons across ranger districts, we converted $s^2\left(\overline{y}_{opt}\right)$ to a coefficient of variation defined as:

$$CV\left(\overline{y}_{opt}\right) = 100 \frac{\sqrt{s^2\left(\overline{y}_{opt}\right)}}{\overline{y}}. \tag{24}$$

Sample size curves for each ranger district were developed by calculating $CV\left(\overline{y}_{opt}\right)$ over a range of sample sizes $(n = 25$ to $n = 400)$ and plotting the results (fig. I). The curves for each ranger district within an ecoregion were similar. This is expected, because within an ecoregion the $CV\left(\overline{y}_{opt}\right)$'s differ only in the stratum weights $(W_h's)$ and mean estimates

Table 5-Individual ranger district DEFFs and RHEs based on all 20 strata (5 site types, 2 seasons, and 2 day types) jointly and then individually by site type, season, and day type

	Strata							
	All 20		5 Site types		2 Seasons		2 Day types	
Ecoregion	DEFF	RHE	DEFF	RHE	DEFF	RHE	DEFF	RHE
1	0.88	0.53	0.93	0.85	0.99	0.98	1.04	1.00
2	.89	S O	1.06	.90	.98	.95	1.02	.98
	1.50	.41	.59	.41	.91	.85	.97	.92
	.51	.19	.50	.40	1.05	.58	.95	.95
3	1.24	.42	1.18	.75	1.06	.91	1.05	.80
	1.11	.47	.98	.67	.98	.92	1.01	.99
	.86	.28	.90	.52	1.00	.77	.89	.76
	1.64	.65	1.56	.97	1.59	.94	1.10	1.09
4	1.08	.14	1.09	.31	.96	.54	1.05	.78
5	.91	.42	.93	.73	.96	.93	.98	.97
6	.90	.21	.96	.56	1.02	1.02	.88	.78
	1.14	.32	1.11	.62	1.02	.70	1.03	.73
	.89	.57	.81	.76	.96	.88	1.02	1.01
	1.07	.43	1.05	.61	.96	.80	.90	.70
8	.98	.57	1.26	.94	1.05	1.02	1.10	.98
	.74	.28	.71	.53	.99	.71	.86	.60
	.63	.18	.76	.40	1.02	.77	1.09	.90
9	.96	.53	.96	.73	1.00	1.00	1.03	.98
	.55	.18	.64	.40	.92	.66	1.06	1.00
10	.82	.44	.86	.71	.91	.82	.96	.91
11	.88	.27	.87	.57	.97	.69	1.00	1.00
	1.34	.72	1.11	.73	1.02	1.02	.99	.98
12	.52	.38	.85	.67	.96	.90	.91	.83
13	1.33	.75	1.21	.83	.94	.94	.95	.94
	2.02	.58	I.26	.92	1.28	.94	.81	.79
14	.98	.47	.96	.61	.95	.73	1.01	.98
	.57	.12	.88	.34	.89	.76	.94	.94
15	1.14	.32	1.04	.64	.97	.93	1.02	.83
16	.83	.20	1.04	.42	.89	.77	.79	.59
	1.41	.40	.99	.60	I .00	.65	.96	.91
Mean	1.01	.40	.97	.64	1.01	.84	.98	.89

DEFF = the design effect; RHE = the relative hypothetical efficiency.

Table 6-Mean relative coeffkients of variation for the strata in ecoregions 2, 6, 13, and 14

Site type	Season[a]	Day type[b]	Ecoregion 2	6	13	14
DUDS	H	H	81	539	59	130
	H	L	1,194	140	106	19
	L	H	I	0	8	41
	L	L	61	59	35	43
GRDS	H	H	349	406	209	143
	H	L	98	89	121	234
	L	H	34	34	61	31
	L	L	168	128	227	35
OUDS	H	H	86	207	708	105
	H	L	77	100	65	57
	L	H	64	50	85	244
	L	L	35	32	74	15
URDS	H	H	225	202	0	103
	H	L	12	126	5	12
	L	H	7	21	122	109
	L	L	67	418	18	6
WBAS	H	H	191	345	279	518
	H	L	109	46	133	551
	L	H	10	93	105	110
	L	L	33	898	42	35

DUDS = day-use developed sites stratum; GRDS = general roaded dispersed sites stratum; OUDS = overnight-use developed sites stratum; URDS = unroaded dispersed sites stratum; WBAS = water-based area sites stratum.

[a] H = high season stratum; L = low season stratum.

[b] H = high (weekends/holidays) day-type stratum; L = low (weekdays) day-type stratum.

$(\bar{y})$ However, the curves also are very similar across ecoregions; specifically, $CV(\bar{y}_{opt})$ was approximately 30 percent when $n = 25$, and between 5 and 10 percent when $n = 400$. Despite large variability in the mean RCV, these sample size curves are quite stable. Generally, large gains in precision [decreasing $CV(\bar{y}_{opt})$] are obtained as n increases to 100; however, beyond that it becomes increasingly difficult to reduce the coefficient of variation.

These curves assume optimum allocation and require fewer observations than proportional allocation for the same coefficient of variation. Comparing our study's average coefficient of variation of 29.3 percent under proportional allocation $(n = 68)$ with the 18 percent (fig. 1) obtained under optimum allocation $(n = 68)$ reveals a reduction of about 11 percent in variability.

To obtain sample size recommendations for a ranger district, it is best to perform a specific pilot study at that ranger district with sample sizes large enough to ensure good variance estimates. We suggest that at least 5 site days be taken in each stratum. However, with 20 strata, a sample size of 100 site days may be too costly. An alternative is to use the mean RCV from table 6 from a comparable ecoregion. If the strata sizes $(N_h\text{'s})$ are known and a rough estimate of $\bar{y}$ is assumed, then $CV(\bar{y}_{opt})$ could be computed. If these quantities are not known, as is usually the case, the relationships shown in figure l will determine the appropriate
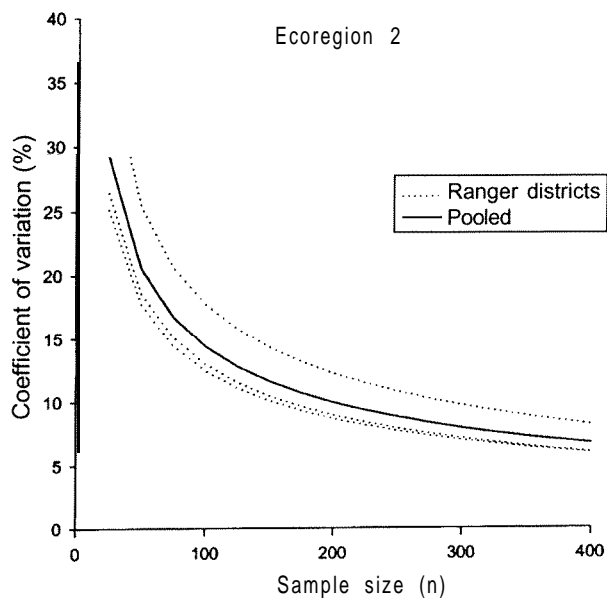
15
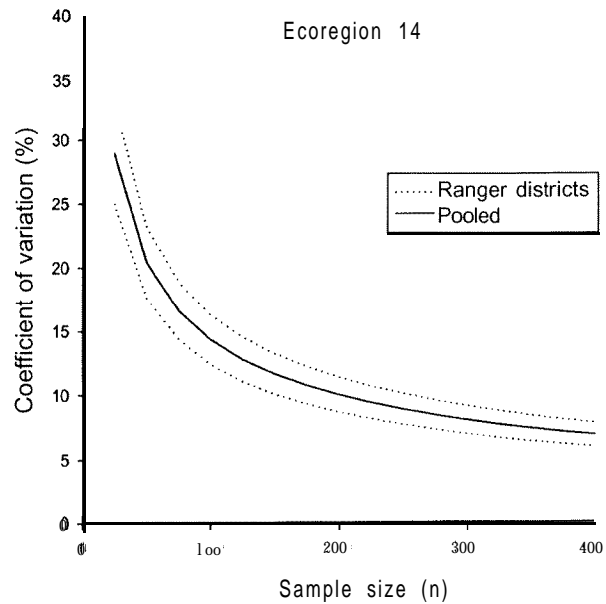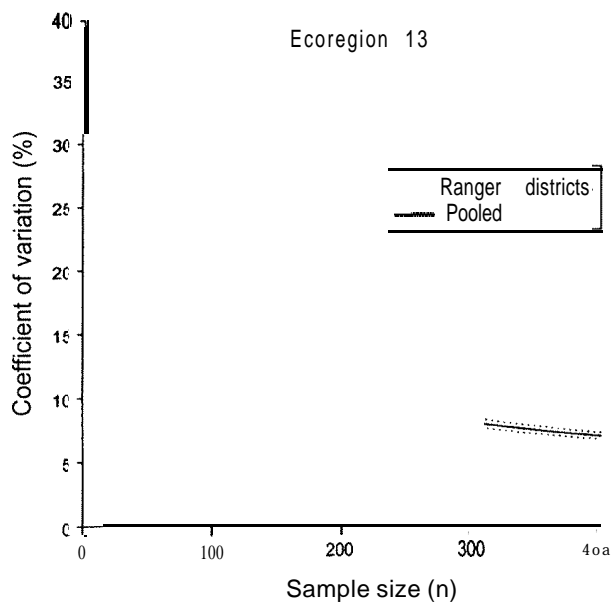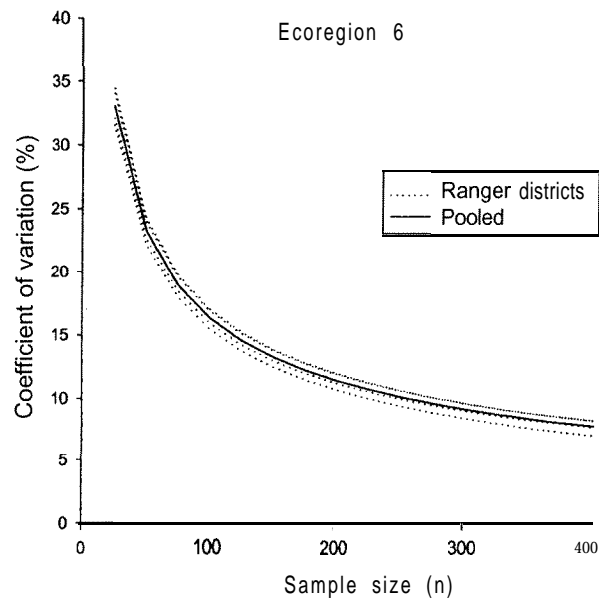
Figure 1--Sample size curves on an individual ranger district basis and a pooled ecoregion basis assuming stratified random sampling under optimum allocation.

sample size for a desired **coefficient** of variation. For example, if a **coefficient** of variation of IO percent is desired, approximately 200 to 250 site days are required under optimum allocation.

Sample size-ecoregion level-Development of sample size recommendations at the ecoregion level is somewhat more complex than at the ranger district level. This is because ecoregion level estimates were based on a three-stage sampling design. For this sampling design under simple random sampling, the sample variance of the estimated mean at the ecoregion level is:

$$s^2(\bar{y}') = \left(1 - \frac{n}{N}\right)\frac{s_1^2}{n} + \frac{n}{N}\left(1 - \frac{m}{M}\right)\frac{s_2^2}{nm} + \frac{n}{N}\left(\frac{m}{M}\right)\left(1 - \frac{k}{K}\right)\frac{s_3^2}{nmk}, \quad (25)$$

where

$N, M,$ and $K \approx$ the total number of primary, secondary, and tertiary stage units in the population;
$n, m,$ and k = the number of these sampled; and
$s_1^2, s_2^2,$ and s_3^2 = the sample variances, respectively.
To develop sample size recommendations from the pilot study, we made several simplifications.

The first simplification assumed that the variance s_3^2 associated with the site-visit estimate on a given site day (third term on the right) was zero. Generally, the site-visit estimate was $y^{(1)}$, which is a function of a supposedly known census quantity, (C), and two estimated quantities, P and $\bar{X}$. Estimation of s_3^2 is complex because $C, P,$ and $\bar{X}$ contribute in a nonadditive manner to the variance. Moreover, sometimes the other site-visit estimators $y^{(2)}$, $y^{(3)}$, and $y^{(4)}$ were used instead, adding more complexity. In addition, the finite population correction $(1-k/K)$ is unknown because K, the total number of last-exiting recreation vehicles from a ranger district, is never observed. Further, the finite population correction for the ranger districts is $n / N \approx 301606 = 0.050$, and for site days it is $m/M \approx 68/8539 = 0.008$, yielding $(n/N)(m/M) \approx 0.0004$, which is small enough that the third component of $s^2(\bar{y}')$ should also be small. Because of that and the difficulty in estimating s_3^2, this term was assumed to be zero in the variance [equation (25)].

The second-stage component [second term to the right of the equal sign in equation (25)] was simplified by assuming the finite population correction $(1 - m/M)$ was 1 .O because $m / A4 \approx 0.008$. Assuming this approximation allowed us to avoid the complexities of an unequal number of site days for various ranger districts. This tends to make the sample size analysis a little more conservative than that obtained by

using the exact value because it results in a slightly larger estimate of $s^2(\bar{y}')$ and, consequently, a higher recommended sample size. Moreover, simple random sampling was assumed instead of stratification because (1) there was considerable concern that strata variances were poorly estimated, which would consequently lead to poor estimates of $s^2(\bar{y}')$ and suboptimal allocation, and (2) this assumption keeps the recommendations on the conservative side.

Variance at the primary stage, s_1^2, was problematic because only nine ecoregions had two or more ranger districts to permit an estimate of the variance. Hence, variance estimates were analyzed at the ecoregion level only for the nine ecoregions where they could be calculated. Although sample size curves could have been computed for the nine ecoregions, we used only the same four ecoregions (2, 6, 13, and 14) used in the previous section to illustrate the concepts.

To develop sample size curves for an ecoregion, we performed a random-effects analysis of variance (Milliken and Johnson 1984, Montgomery 1976) using **Proc** VARCOMP (SAS Institute Inc. 1989). Results included the typical variance components between ranger districts, σ_1^2, and within ranger districts, σ_2^2, which were used to estimate $s^2 = \sigma_1^2 + \sigma_2^2/\bar{m}$ and $s_2^2 = \sigma_2^2$ (Marcuse 1949, Zamoch and others 1993) shown in table 7. Note that $\bar{m}$ is an unequal sample size mean number of site days per ranger district for the given ecoregion computed according to Montgomery (1976, p. 53) and Milliken and Johnson (1984, p. 2 19). Using our simplifications, the known N and M values (number of primary and secondary stage units) and estimates of s_1^2 and s_2^2, $s^2(\bar{y}')$ [equation (25)] was computed on an ecoregion basis for a range of n (2, 4, 8, and 16 sampled ranger districts) and m (10 to 100 site days) and converted into coefficients of variation. Figure 2 illustrates general trends in these coefficients of variation for four of the ecoregions. The general result is that the coefficient of variation is improved little when m is greater than 40. However, increasing n diminishes the coefficient of variation substantially. For instance, the initial gain in precision by doubling n from 2 to 4 is almost as much as what is gained by quadrupling from 4 to 16.

Although the coefficient of variation will decrease as n and m increase, a coefficient of variation of < 15 percent is probably impractical to obtain unless more than eight ranger districts are sampled in each ecoregion. Generally, it is better to increase n and decrease m for a given total sample size (nm). The dotted line in figure 2 shows the $nm = 200$ line, which is the coefficient of variation when a total of 200 site

Table 7-Variances s_1^2 and s_2^2 used to compute the variance of the mean site-visit estimate at the ecoregion level assuming two-stage sampling with a simple random sample at both stages

Ecoregion	N	n	s_1^2	s_2^1	Mean
2	57	3	1,891	65,988	98.8
3	82	4	2,006	33,645	61.1
6	107	4	202	17,668	27.4
8	60	3	995	19,104	56.8
9	35	2	2,193	139,534	172.9
11	37	2	2,336	7,448	40.6
13	29	2	60	1,169	17.1
14	35	2	968	20,490	50.8
16	41	2	8,141	66,376	76.4

N = number of ranger districts in the ecoregion; n = number of ranger districts sampled in the ecoregion.

days are sampled with n ranger districts and m site days per ranger district. Drastic reductions to the coefficient of variation are possible for a fixed nm by increasing n and decreasing m. Such gains in precision at a fixed $nm = 200$ are obtainable if there is little additional cost or a fixed overhead cost for sampling more primary units (ranger districts). However, if these costs are a function of the number of primary units, for example, due to increased travel costs, then cost must be considered, and the nm line should be viewed with caution. In any event, it should be reassuring to the user of figure 2 that the relationships are similar among the four ecoregions.

Sample size-national level-Results from an ecoregion analysis easily can be extended to the national level by specifying a common coefficient of variation for all ecoregions. The sample size curves (fig. 2) can be used to obtain an approximate n and m for all ecoregions based on the common coefficient of variation. If this is done, then the level of precision achieved when the national estimate is obtained by combining the ecoregions [equation (16)] is at least at this level. This is easily shown as follows. Let there be two ecoregions where it is specified that the coefficient of variation of both is P percent. Thus,

$$CV_i = 100 \frac{\sqrt{s_i^2}}{T_i} = P ,$$

(26)

for all i

where

T_i = the total estimate for ecoregion i, and
s_i^2 = its variance.
Solving for s_i^2 yields

$$s_i^2 = \frac{P^2 T_i^2}{100^2} ,$$

(27)

Because the total estimate is simply $T = T_1 + T_2$, the coefficient of variation relationship becomes

$$CV(T) = 100 \frac{\sqrt{s_1^2 + s_2^2}}{T_1 + T_2} = 100 \frac{\sqrt{\frac{P^2 T_1^2}{100^2} + \frac{P^2 T_2^2}{100^2}}}{T_1 + T_2}$$

$$= 100 \frac{P}{100} \frac{\sqrt{T_1^2 + T_2^2}}{T_1 + T_2} \leq P ,$$

(28)

because

$$\sqrt{T_1^2 + T_2^2} \leq T_1 + T_2 ,$$

(29)

for T_1 and $T_2 \geq 0$. The generalization to more than two ecoregions is obvious.

Sample size-alternative specification-The sample size recommendations given above are based on sampling with a desirable coefficient of variation. However, the sample size issue is often stated in terms of being within an allowable percent error (D) with a specified probability. To see the relationship between these two specifications, note that a coefficient of variation of 20 percent implies that one standard error is 20 percent of the mean. Assuming the normal distribution, one is 68 percent confident that the mean is within $D = 20$ percent of the true mean. This is the basis underlying figure 2. To design a survey with other, more typically used probability levels, one can simply redefine the coefficient of variation as the desired percent error D divided by the appropriate z-value from the standard normal distribution (Montgomery 1976). For instance, if a survey is to be designed where the specified allowable error is $D = 20$ percent at the 90-percent probability level, then redefine the coefficient of variation to be $CV = 20/1.645 = 12.16$ percent where 1.645 is the z-value that corresponds to the 90-percent probability level. Then select an appropriate ecoregion sample size curve from figure 2 and draw a horizontal line at the coefficient of variation = 12.16 position. This line will represent a set of n and m values from which the number of ranger districts and site days per ranger district, respectively, are determined to ensure being within $D = 20$ percent error at the 90-percent probability level.
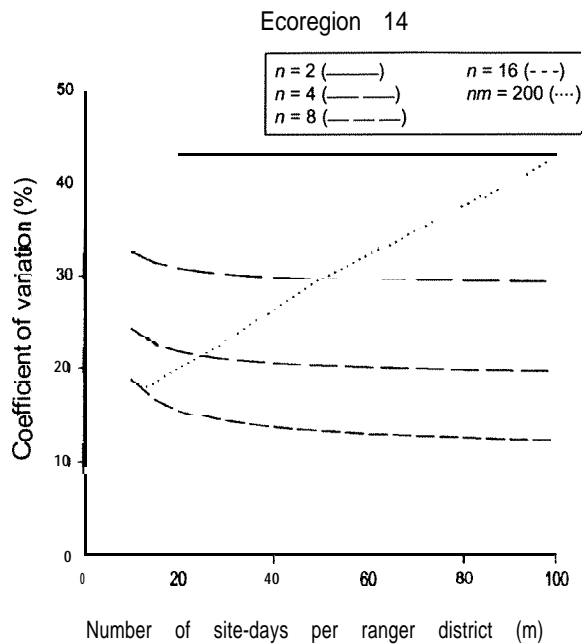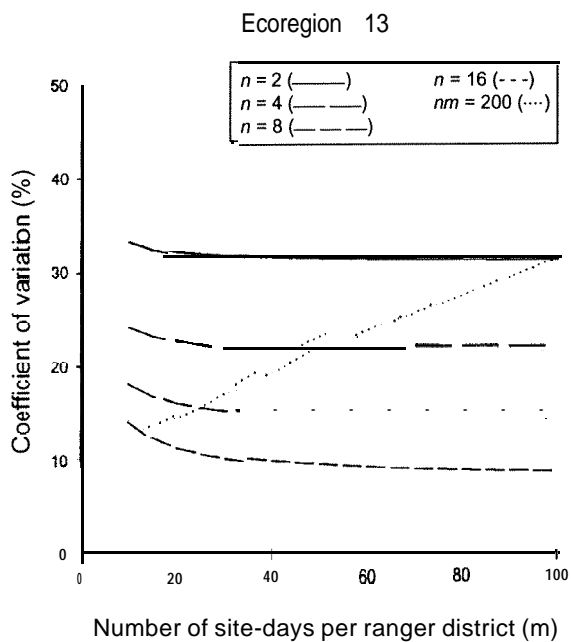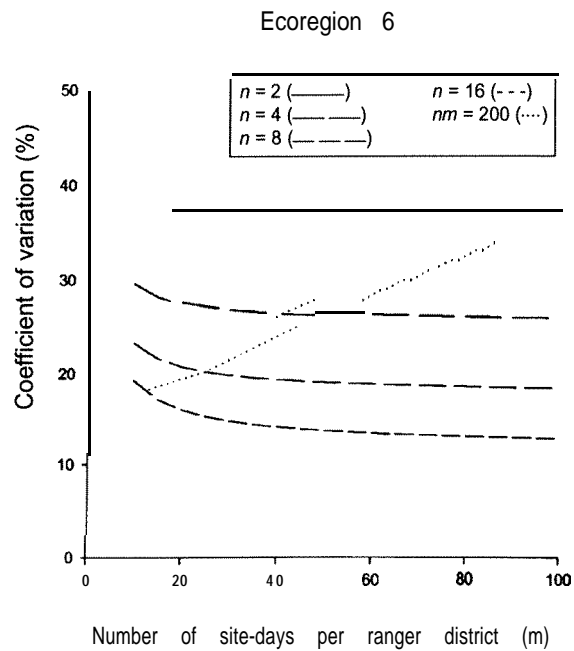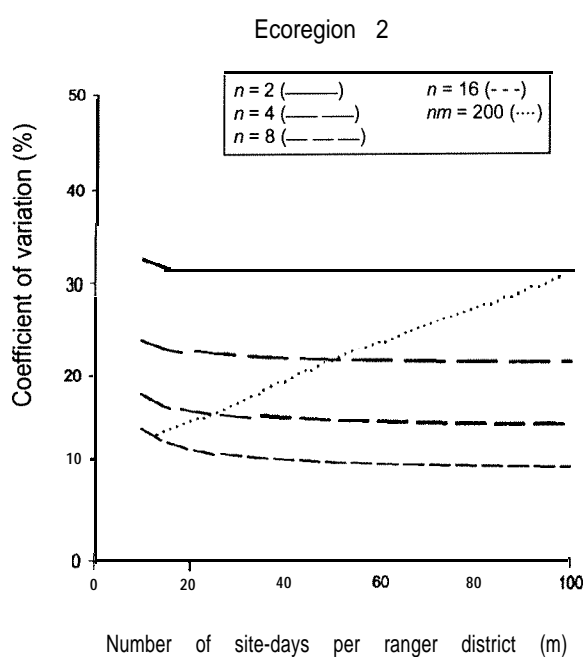
Figure 2-Coefficients of variation for four of the ecoregions for estimating site visits assuming a two-stage sampling design where n = the number of sampled ranger districts per ecoregion and m = the number of sampled site days per ranger district.

Conclusion

This study is the first statistically based, outdoor recreation survey performed on the entire National Forest System. Among the ranger districts sampled, daily site-visit estimates varied considerably and averaged 6.1, and total annual site visits averaged 536,378 per ranger district. The variability of these estimates also ranged widely with a coefficient of variation of 10.4 to 66.9 percent, averaging 29.3 percent, indicating the potential for reasonably good estimates for some ranger districts, but also some difficulty for others. The total annual site-visit estimate for the entire National Forest System was 330,772,894 with a coefficient of variation of 20 percent.

Due to great variability and small sample size, it was difficult to evaluate stratification variables. Generally, based on the analysis of variance, DEFF, and RHE, site type appeared to be the most important; and it should be incorporated in future surveys. Season and day type appeared somewhat less important, possibly because the variance reduction that they possessed jointly was partitioned individually to these two variables. For that reason, we suggest that season and day type should be combined into one use-level stratification variable with only two or three levels. There were too few ranger districts in the ecoregion stratification to evaluate ecoregion stratification.

The sample size curves presented should give an estimate of the sample size requirements for specified levels of precision based on the coefficient of variation. Generally, site visits at the ranger district level could be estimated with a coefficient of variation of 15 percent by sampling 100 site days using a stratified random sampling design with optimum allocation. The sample size curves at the ecoregion level present various combinations of number of ranger districts and site days per ranger district required to achieve a specified level of precision using a two-stage sampling design. These sample size guidelines can be used in planning national level outdoor recreation surveys.

Literature Cited

Bailey, R.G. 1995. Description of the ecoregions of the United States. Misc. Publ. 1391. Washinton, DC: U.S. Department of Agriculture, Forest Service. 108 p. + map.

Bailey, R.G.; Avers, P.E.; King, T.; McNab, W.H., eds. 1994. Ecoregions and subregions of the United States (map). Washington, DC: U.S. Department of the Interior, Geological Survey. 1:7,500,000; colored. [Accompanied by a supplementary table of map unit descriptions compiled and edited by W. Henry McNab and Robert G. Bailey. Prepared for the U.S. Department of Agriculture, Forest Service.1

Cochran, W.G. 1977. Sampling techniques. 3rd ed. New York: John Wiley. 428 p.

Conover, W.J. 1980. Practical nonparametric statistics. 2" ed. New York: John Wiley. 493 p.

Gregoire, T.G.; Buhyoff, C.J. 1999. Sampling and estimating recreational use. Gen. Tech. Rep. 456. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 39 p.

James, G.A. 1967. Recreation use estimation on Forest Service lands in the United States. Res. Note SE-79. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 8 p.

James, G.A.; Henley, R.K. 1968. Sampling procedures for estimating mass and dispersed recreation use on large areas. Res. Pap. SE--31 Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 15 p.

James, G.A.; Ripley, T.H. 1963. Instructions for using traffic counters to estimate recreation visits and use. Res. Pap. SE-3. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 12 p.

Kish, L. 1965. Survey sampling. New York: John Wiley. 643 p

Kocis, S.M. 1986. The adoption and diffusion of methods for estimating recreation use in the Pacific Northwest Region of the U.S. Forest Service. Moscow, ID: University of Idaho. 161 p. M.S. thesis.

Marcuse, S. 1949. Optimum allocation and variance components in nested sampling with an application to chemical analysis. Biometrics. 5: 189–206.

Milliken, G.A.; Johnson, D.E. 1984. Analysis of messy data. Volume I: Designed experiments. New York: Van Nostrand Reinhold Co. 473 p.

Montgomery, D.C. 1976. Design and analysis of experiments. New York: John Wiley. 418 p.

Rao, J.N.K. 1962. On the estimation of the relative efficiency of sampling procedures. Annals of the Institute of Statistical Mathematics. 14: 143–150.

Ruark, C.A.; Zarnoch, S.J. 1992. Soil carbon, nitrogen, and fine root biomass sampling in a pine stand. Soil Science Society of America Journal. 56(6): 1945- 1950.

SAS Institute Inc. 1989. SAS/STAT user's guide. Version 6. 4'" cd. Cary, NC: SAS Institute Inc. 846 p. Vol. 2.

SAS Institute Inc. 1990. SAS procedures guide. Version 6. 3'" ed. Cary, NC: SAS Institute Inc. 705 p.

Zarnoch, S.J.; Marx, D.H.; Ruehle, J.L.; Baldwin, V.C. 1993 Sampling open-top chambers and plantations for live fine-root biomass of loblolly pine. Res. Note SE-368. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 19 p.

A pilot sampling design is described for estimating site visits to National Forest System lands. The three-stage sampling design consisted of national forest ranger districts, site days within ranger districts, and last-exiting recreation visitors within site days. Stratification was used at both the primary and secondary stages. Ranger districts were stratified based on Bailey's ecoregions, while site days were stratified based on site type, season, and day type. Statistical methodology is presented to derive site-visit estimates at the site day, ranger district, and national levels. Results are presented to illustrate the magnitude of the site-visit estimates, their variability, and confidence intervals. With such information, an evaluation of the stratification variables is presented using the design effect and the relative hypothetical efficiency. Sample size analysis is performed to provide recommendations for future sample surveys to meet specified levels of precision